The Plant Genome
Original Research

# The Effect of Tropical Sorghum Conversion and Inbred Development on Genome Diversity as Revealed by High-Resolution Genotyping

Robert R. Klein, John E. Mullet, David R. Jordan, Frederick R. Miller, William L. Rooney, Monica A. Menz, Cleve D. Franks, and Patricia E. Klein*

R.R. Klein, USDA-ARS-SPARC, College Station, TX 77845; J.E. Mullet, Dep. of Biochemistry and Biophysics, Texas A&M University, College Station, TX 77843; D.R. Jordan, Dep. of Primary Industries and Fisheries, Hermitage Research Station, Warwick, Queensland 4370, Australia; F.R. Miller, MMR Genetics Inc., Vega, TX 79092; W.L. Rooney, Dep. of Soil and Crop Science, Texas A&M University, College Station, TX 77843; M.A. Menz, Syngenta Seeds SAS, 57 Chemin des Amandiers, 31700 Beauzelle, France; C.D. Franks, USDA-ARS-PSGDR, Lubbock, TX 79415; P.E. Klein, Dep. of Horticulture and Institute for Plant Genomics and Biotechnology, Texas A&M University, College Station, TX 77843.

Received 7 June 2007. *Corresponding author (pklein@tamu.edu).

## Abstract

Graphical genotypes have been generated for a set of sorghum [*Sorghum bicolor* (L.) Moench] germplasm, which includes selected public inbreds, germplasm from the world collection, and ancestral lines central to the early breeding efforts of sorghum. We have focused our present examination on sorghum chromosome SBI-06, which encodes $ma_1$ and $dw_2$, two genes critical to sorghum improvement dating to the original introduction of tropical sorghums into the United States. Utilizing the pedigree relationship between sorghum cultivars, the patterns of genetic variation were detailed within segmental chromosomal blocks of SBI-06. Segmental genomic blocks were traced back through multiple generations of a pedigree, often back to founder tropical accessions. The graphical genotypes reveal genomic signatures of historical breeding decisions, especially evidence of directional selection during the conversion of tropical accessions to temperate adaptation. This information is central to our efforts to understand those crop improvement processes that have shaped the genomic diversity of elite sorghum cultivars.

The chromosomes of any particular genotype are predicted to be a mosaic of ancestral chromosomal segments, and the pattern of genetic variation provides insight into the breeding history of the crop (Jordan et al., 2004; Lorenzen et al., 1995; Pestsova and Roder, 2002; Russell et al., 2000; Shoemaker et al., 1992; Sjakste et al., 2003). Constructing high-resolution graphical genotypes can identify genomic regions with great allelic diversity and other regions, often quite large, which have been fixed due to factors that include domestication bottlenecks, selection pressure, and genetic hitchhiking (Hamblin et al., 2004, 2005, 2006; Wade et al., 2002; Wade and Daly, 2005; Yalcin et al., 2004). This type of combined pedigree–molecular marker analyses can provide a historical perspective of the effect of domestication or selection pressure during cultivar improvement on discrete regions of the genome. Sorghum [*Sorghum bicolor* (L.) Moench] appears well suited for this type of analysis due to the

**Abbreviations:** AFLP, amplified fragment length polymorphism; CMS, cytoplasmic male sterility; Indel, insertion-deletion; IBD, identical by descent; IBS, identical by state; PBMASS, Pedigree Based Marker Assisted Selection System; QTL, quantitative trait loci; SSR, simple sequence repeat.

Published in Crop Sci. 48(S1) S12–S26. Published 8 Feb. 2008.
doi:10.2135/cropsci2007.06.0319tpg

677 S. Segoe Rd., Madison, WI 53711 USA

development of robust genomic resources (Bowers et al., 2003, 2005; Klein et al., 2000, 2003) and the availability of detailed documentation of sorghum pedigrees dating back to its introduction into the United States. The purpose of the present study was to provide a historical prospective of sorghum improvement in the 20th century and detail how breeding decisions have impacted chromosomal haplotype diversity across a panel of cultivars dating from sorghum's introduction into the United States.

Sorghum is not native to the Western Hemisphere, but its introduction into the New World came from Africa by way of slave ships more than 200 years ago (Quinby, 1974, 1975; Rooney and Smith, 2000; Smith and Frederiksen, 2000). There are distinct phases associated with grain sorghum improvement in the United States including (i) the introduction of a limited number of founder cultivars (1874–1908), (ii) the selection for short-stature, early maturing plants from heterogeneous populations (1904–1936), (iii) breeding of improved, combine-harvestable cultivars (1930s–1940s), (iv) the advent of hybrid seed production (1946–present), and (5) the conversion of exotic tropical accessions (tall, photoperiod sensitive) to temperate adaptation (short, photoperiod insensitive) and the use of this diverse germplasm in breeding programs (1963–present). Each of these improvement phases had profound effects on sorghum production in the 20th century across the semiarid regions of the world.

Two tropical varieties, Milo Maize and Guinea Kafir, may be considered the two most critical founder genotypes for the establishment of grain sorghum production in the United States. Both of these cultivars mature too late for the temperate zone (Quinby, 1974; Smith and Frederiksen, 2000). However, farmers and breeders soon selected mutants of short stature (Dwarf Milo, Double Dwarf Milo), early maturation (Early White Milo), and mutants that were both early and short stature (Blackhull Kafir, Double Dwarf White Sooner Milo). Later crosses of Kafirs and Milos produced superior dwarf cultivars with erect panicles, which was critical for the harvesting with combines. Wheatland (circa 1931), a selection from a Kafir × Milo cross, could be harvested with a wheat combine. When additional Kafir × Milo cultivars including Martin, Caprock, Combine Kafir-60, Plainsman, Combine 7078, Redlan, and Redbines were released, the need for combine-harvestable cultivars for sorghum production from Texas to Nebraska was met.

The ability of farmers to select early maturing, short-stature sorghums relates to the existence of defined sets of dwarfing and maturity genes. Classical genetic studies in sorghum identified four maturity loci, designated $Ma_1$, $Ma_2$, $Ma_3$, and $Ma_4$ (Quinby, 1974, 1975; Smith and Frederiksen, 2000), with an allelic series at $ma_1$ and $ma_3$. More recently, two additional maturity genes, designated $Ma_5$ and $Ma_6$, were described that influence photoperiodism and floral initiation in sorghum (Rooney and Aydin, 1999). Dominance at each of the *Ma* loci generally causes late flowering, and of the first four maturity genes ($Ma_1$–$Ma_4$), $Ma_1$ has the largest impact on flowering date. Four dwarfing genes have been identified in sorghum that were designated $Dw_1$ through $Dw_4$ (Quinby, 1974, 1975). The *dw* loci are unlinked, and all dwarfing genes exhibit dominance for tallness. Genetic analyses indicate that growers, during the first 40 years of the 20th century, had selected recessive mutations at the $Ma_1$ to $Ma_3$ and $Dw_1$ to $Dw_3$ loci resulting in early maturing, short-stature cultivars suitable for mechanized harvest.

The selection for photoperiod insensitive, short-stature mutants made sorghum suitable for mechanical harvest, but did not increase grain yields (Quinby, 1974; Smith and Frederiksen, 2000). As early as 1927 (Conner and Karper, 1927), hybrid vigor was recognized in sorghum. Before the discovery of cytoplasmic male sterility (CMS) in sorghum, genetic male sterility in Day Milo was utilized in crosses with Blackhull Kafir, which represents the first generation hybrid dating to the late 1940s (Quinby, 1974). By 1952, it was proven that CMS existed in Dwarf Yellow Milo with nuclear-encoded male sterile genes residing in Texas Blackhull Kafir (Stephens and Holland, 1954). Fertility restorer (*RF*) genes were quickly discovered in Milo, thereby completing the necessary genetic components for commercial production of hybrid sorghum and CMS quickly replaced the use of the Day Milo genetic male sterile system. By 1960, eight years after the discovery of CMS in sorghum, the acreage planted to hybrids reached 95% while grain yields doubled during this same period (Quinby, 1974; Smith and Frederiksen, 2000).

Nearly every facet of grain sorghum improvement in the period from 1900 to 1960 can be directly traced to the introduction of a very limited number of genotypes dominated by Milos and Kafirs. While early genetic gains were impressive, scientists also recognized that genetic bottlenecks had created a rather narrow genetic base, which could limit the potential for genetic progress and leave the crop vulnerable to biotic and abiotic stresses. Armed with knowledge of the inheritance of plant height and maturity in sorghum, in 1963 the Sorghum Conversion Program was established to convert many of the tropical accessions to temperate adaptation (Quinby, 1974; Rosenow and Dahlberg, 2000; Stephens et al., 1967). The conversion program was designed to

move recessive dwarfing and photoperiod insensitive genes from a four-dwarf temperate zone variety into the genomes of exotic lines. Through this program, over 840 converted and partially converted lines have been developed thereby providing new, diverse germplasm that now provides an important source of the germplasm used in sorghum improvement programs throughout the world.

The development of genomic resources for sorghum, including saturated genetic maps (Bowers et al., 2003; Menz et al., 2002), permits a detailed historical examination of the effect of the aforementioned distinct phases associated with grain sorghum improvement on the genetic diversity for defined chromosomal regions. Of particular interest is sorghum chromosome SBI-06 that encodes $ma_1$ and $dw_2$ (Lin et al., 1995), two loci that were critical to sorghum improvement dating back to the original introductions of tropical Milo and Kafir. In the present study, graphical genotypes were constructed for a series of public sorghum lines on the basis of their seminal role in grain sorghum production throughout the 20th century. Included in this study are modern public inbreds, earlier generation cultivars and plant introductions dating to the turn of the 20th century. Genotyping of founder introductions (e.g., Kafir, Milo) permitted specific haplotype blocks to be traced back through the pedigrees of elite cultivars to sorghum's introduction into the United States. Finally, we examine cultivars arising from the Sorghum Conversion Program to determine the effectiveness of this breeding scheme in recovering the exotic haplotype of chromosome SBI-06 and thereby decreasing the man-made bottlenecks created during sorghum introduction and early germplasm improvement efforts.

# Materials and Methods

## Plant Material

Fifty sorghum inbreds previously examined (Menz et al., 2004) were included in this study. In brief, these lines represent public materials that are critical to sorghum improvement dating to early selections from Milo and Kafir introductions. Cultivars represent 13 A/B-lines and 18 R-lines present in the pedigrees of many hybrid sorghums. Also included were 16 lines out of the Sorghum Conversion Program widely used as progenitors of elite inbreds. A series of ancestral cultivars were also genotyped in the present study, dating from the introductions of Standard Yellow Milo (circa 1900) and Blackhull Kafir (circa 1880), and early selections including Early White Milo, Dwarf Yellow Milo, Double Dwarf Yellow Milo, Day Milo, and Double Dwarf White Sooner Milo. Two maturity genotypes, 100M and SM100, were included as reference genotypes for $Ma_1$ (100M) and $ma_1$ (SM100). Lastly, included in this study were a series of tropical accessions from which converted lines were derived.

## Linkage Map Analyses of $ma_1$ and $dw_2$ Loci

An $F_2$ population (110 individuals) derived from a cross of BTx406 ($ma_1$, $Ma_2$, $Ma_3$, $Ma_4$) and 100M ($Ma_1$, $Ma_2$, $Ma_3$, $Ma_4$) was created for linkage analysis of $ma_1$. Phenotypic classification of floral initiation was conducted as previously described (Rooney and Aydin, 1999). Seed from each late maturing $F_2$ plant was progeny tested either in the field or the greenhouse to differentiate homozygous ($Ma_1$, $Ma_1$) and heterozygous ($Ma_1$, $ma_1$) individuals. A series of 45 amplified fragment length polymorphism (AFLP) primer combinations (+3/+3 selectivity) was initially used for linkage analysis. As this population segregates for only one major maturity locus, plant maturity was included in linkage analyses as a morphological-trait locus. A regional map was constructed with approximately 360 AFLP markers using MAPMAKER/EXP version 3.0 as previously detailed (Klein et al., 2001). Results from this linkage analysis were in agreement with previous research that placed $ma_1$ on sorghum chromosome SBI-06 (Lin et al., 1995; Ulanch, 1999). Additional genetic markers spanning the $ma_1$ locus were identified by aligning the regional map to the high-density genome map of sorghum (Menz et al., 2002). The location of the $ma_1$ locus was refined by identifying crossover events within the delimited locus, with the position of these events being placed on the high-density genetic map of sorghum (Menz et al., 2002; http://sorgblast3.tamu.edu/index.html; verified 7 Jan. 2008).

An $F_2$ population (85 individuals) derived from a cross of BTx3197 ($dw_1$, $Dw_2$, $dw_3$, $dw_4$) and BTx616 ($dw_1$, $dw_2$, $dw_3$, $dw_4$) was utilized for mapping $dw_2$. Genetic studies (Quinby, 1974, 1975) have shown that $Ma_1$ and $dw_2$ are linked, which is in agreement with the map location as determined previously (Lin et al., 1995). Based on these observations, markers spanning the region of SBI-06 harboring $ma_1$ were utilized to localize $dw_2$ on the high-density genetic map of sorghum. Plant height to the base of the flag leaf was used for phenotypic classification. All other mapping activities and refinement of the trait loci were as detailed above for $ma_1$.

## Germplasm and DNA Isolation

Seeds from inbreds were obtained from public breeders and the ARS Plant Genetic Resources Conservation Unit (Griffin, GA). Where possible, multiple samples of a given inbred were obtained from different sources to minimize the possibility of improperly

identified seed stocks. DNA was isolated from seedling leaf tissue as previously described (Menz et al., 2002). DNA was isolated from ~12 seedlings to construct a representative sample of individuals from each seed stock.

## Genotypic and Data Checking

A series of ~100 mapped, randomly spaced markers spanning sorghum SBI-06 were utilized for genotyping the 69 cultivars examined in this study. Markers included simple sequence repeats (SSRs), dominant and codominant AFLPs, and insertion-deletions (Indels). Markers with a rare allele frequency of ≤5% were examined for scoring errors, and markers that were difficult to score or with numerous missing observations were eliminated. The result was a set of 87 informative markers spanning SBI-06. Information on published markers used in this study can be found at http://sorgblast3.tamu.edu/index.html.

Amplified fragment length polymorphism markers were analyzed on dual dye LI-COR 4200 IR$^2$ gel detection systems (LI-COR, Lincoln, NE) as previously detailed (Klein et al., 2000; Menz et al., 2002). Amplified fragment length polymorphism size estimates were calculated with Bionumerics software version 2.5 (Applied Maths BVBA, Kortrijk, Belgium). Simple sequence repeats and Indels were analyzed on an ABI 3130 Genetic Analyzer (Applied Biosystems, Foster City, CA) with alleles called with GeneMapper Software v4.0 (Applied Biosystems). To reduce the genotyping error rate to a minimal level, marker data were scanned for the appearance of rare variant alleles. As even a modest genotyping error rate gives the appearance of rare variant haplotype blocks that do not exist in nature, rare variant alleles were closely examined for discrepancies tracing to errors in data collection or analysis. Excel spreadsheets containing genotypic observations for all cultivars can be found in the supplemental material associated with this manuscript.

## Pedigree Information

Pedigree information for each inbred included in this study was obtained by the following: consultation with scientists integrally involved with sorghum development over the last 50 years; from published literature; and from pedigree notes maintained by F.R. Miller (MMR Genetics). A Microsoft Office Excel 2003 spreadsheet (Microsoft Corp., Redmond, WA) containing these extended pedigrees can be found in the supplemental material associated with this manuscript.

## Haplotype Blocks and Ancestral Identity

The Pedigree Based Marker Assisted Selection System (PBMASS) software developed by the Department of Primary Industries, Queensland, Australia, was utilized for haplotype block partitioning and assigning identities by descent for each haplotype block. Details of the methodologies implemented in PBMASS are described elsewhere (Jordan et al., 2004). In brief, PBMASS provides graphical representations of haplotype blocks that are identical by descent (IBD) or identical by state (IBS). Identical by descent calculations in PBMASS utilize actual marker data combined with values inferred from the information in the extended pedigrees and the existing marker data. The density of haplotype blocks is improved by applying flanking marker inference of IBD to loci for which the IBD calculations were not successful. The default settings of PBMASS were used in the present study to infer IBD regions of sorghum SBI-06. IBD haplotype blocks were assigned a specified color within PBMASS. To confirm the IBD haplotype blocks of PBMASS and identify haplotype blocks that are IBS, marker data were further analyzed with Excel's advanced filter tool. IBS determinations are the consequence of one (or more) parental lines not being examined in this study, which therefore precludes an accurate IBD determination. To identify variants of a given haplotype block (differ in one marker observation), the advanced filter tool was employed in conjunction with the wildcard option. Each marker observation in a haplotype block was replaced sequentially with the wildcard character in Excel, thereby identifying haplotype blocks that differ by a single marker entry. Before assigning a new haplotype block identity, each variant marker block was examined for possible genotyping errors that can result in unique, spurious haplotype blocks.

## Graphical Genotypes

Graphical depiction of IBD haplotype blocks initially utilized the PBMASS software package. These results were utilized as template to create the graphical genotypes in Adobe Illustrator CS2 (Adobe Systems, San Jose, CA). Subsequently, IBS haplotype blocks delimited in Excel (detailed above) were positioned on graphical genotype scaffolds created in Adobe Illustrator. Graphical genotypes, therefore, represent a graphical depiction of both IBD and IBS relationships of the sorghum cultivars. Base colors represent distinct ancestral "founder" genotypes (e.g., Milo, Kafir) or the different sorghum races. Different tones of these colors represent variant haplotype blocks of founder genotypes caused by recent mutation or by intragenotype variation that existed in ancestral populations before the directed inbreeding part of cultivar creation. Additional base colors are assigned to the haplotypes of inbreds when first encountered

in pedigree analyses. Diagonal lines indicate blocks that are IBS with like-colored blocks. Uncolored blocks denote extended strings of uninformative monomorphic markers within ancestors of a pedigree, or unique blocks where IBS was not predicted by Excel's advanced filter tool. Haplotype blocks that are inconsistent with the reported pedigree of the given cultivar are marked with cross-hatched lines.

# Results

## Marker Analysis

A series of 87 markers spanning sorghum chromosome SBI-06 were utilized in the present study. SBI-06 markers utilized consisted of 34 SSRs (prefixes *Xtxp, TS, CS, Xcup*, *SDB*), 52 AFLPs (prefix *Xtxa*, 11 multiallelic), and one Indel (*Xtxi20*). While markers were not evenly spaced, all regions of the linkage group were well covered, averaging one marker per 1.4 cM. To avoid spurious linkage, only markers that could be placed at a high LOD score were utilized, with a low of LOD 10.3 (*Xtxp6*) and a modal LOD score of 33.0 for all examined markers. The average allelic frequency of all markers utilized was ~3.5 alleles per locus, with a low being displayed by AFLPs (2.2 alleles per locus), while SSRs averaged 5.4 alleles per locus. The marker set chosen provided a level of marker redundancy, thereby increasing the likelihood of differentiating any two genotypes at any given region of the linkage group. In addition, the physical order of most of the markers along the linkage group have been determined based on marker integration into the sorghum physical map or by identifying marker sequences (SSRs, Indels) within the genome sequence of sorghum (http://www.phytozome.net/sorghum; verified 7 Jan. 2008).

## Mapping $ma_1$ and $dw_2$

Lin et al. (1995) have previously mapped the $ma_1$ and $dw_2$ loci in sorghum. By cross-referencing the sorghum maps of Menz et al. (2002) and Bowers et al. (2003), $ma_1$ and $dw_2$ were placed on the distal end of SBI-06. Since the map locations of the $ma_1$ and $dw_2$ loci were critical to the present study, linkage analyses of these genes were re-examined in populations that either segregate for a single major maturity ($ma_1$) or height gene ($dw_2$). The results of these mapping efforts are displayed in Fig. 1 (left panel).

In mapping $ma_1$, a total of five genetic markers on the distal end of SBI-06 were polymorphic in the cross of 100M ($Ma_1$, $Ma_2$, $Ma_3$, $Ma_4$) and BTx406 ($ma_1$, $Ma_2$, $Ma_3$, $Ma_4$). Four of the five polymorphic markers (*Xtxa4001*, *Xtxa2321*, *Xtxa3550*, *Xtxi20*) clustered between ~11 to 21 cM on SBI-06, suggesting that the $ma_1$ locus resides near this region. Examination of informative $F_2$ individuals (e.g., individuals displaying crossover events within the trait locus) implicate the genomic region immediately downstream of *Xtxi20* as harboring the $ma_1$ locus (data not shown). Using a similar mapping strategy, $dw_2$ was mapped in a cross of a three-dwarf (BTx3197; $dw_1$, $Dw_2$, $dw_3$, $dw_4$) and a four-dwarf inbred (BTx616; $dw_1$, $dw_2$, $dw_3$, $dw_4$). Mapping plant height within this $F_2$ population localized the $dw_2$ locus to a region adjacent to $ma_1$, with the locus delimited by markers *Xtxa2124* (21.1 cM) and *Xtxa3407* (29.2 cM). Further crossover events were not detected in either of these populations, which precluded additional refinement of the loci at the time of this paper.

Supportive evidence for the map location of $ma_1$ was found by examining the SBI-06 haplotypes of maturity Genotypes 100M and SM100 (Fig. 1, center panel). Both 100M and SM100 are full sibs derived from a cross of Double Dwarf Yellow Milo and Early White Milo (Quinby, 1974, 1975). SM100 is a temperate genotype ($ma_1$, $Ma_2$, $Ma_3$, $Ma_4$) while 100M is essentially tropical ($Ma_1$, $Ma_2$, $Ma_3$, $Ma_4$), with the two lines differing only at the $ma_1$ locus. The SBI-06 haplotypes of 100M and SM100 are IBD except for a single haplotype block that coincides with the proposed map location of $ma_1$. Pedigree analysis shows that this $ma_1$–spanning haplotype block of SM100 is IBD with Early White Milo ($ma_1$) while the haplotype block of 100M is IBD with Double Dwarf Yellow Milo ($Ma_1$).

The graphical genotypes of BTx406 ($ma_1$, $dw_2$) and the genotypes of its ancestors provide further support for the proposed location of the $ma_1$ and $dw_2$ loci, while also permitting $ma_1$ and $dw_2$ to be traced back through its pedigree to a founder genome (Fig. 1, right panel). BTx406, in most cases, was the donor line used in the Sorghum Conversion Program (Quinby, 1974; Rosenow and Dahlberg, 2000; Stephens et al., 1967) where $ma_1$ and $dw_2$ were introduced into the background of tall, photoperiod-sensitive exotic sorghum. The combined pedigree-graphical genotypic analysis revealed that BTx406 acquired $ma_1$ from Early White Milo and $dw_2$ from Double Dwarf Yellow Milo. In detail, graphical genotypes showed that the haplotype block spanning $ma_1$ (yellow block, labeled "r") can be traced through SA403, to Double Dwarf White Sooner Milo (or Day Milo), and finally to the founder genome of Early White Milo. Similarly, the region spanning $dw_2$ (red block, labeled "r") can be traced through SA403, to Double Dwarf White Sooner Milo, and back to the founder cultivar Double Dwarf Yellow Milo. These results are in agreement with the graphical genotypes of SM100 and 100M (Fig. 1, center panel), while also illustrating Quinby's

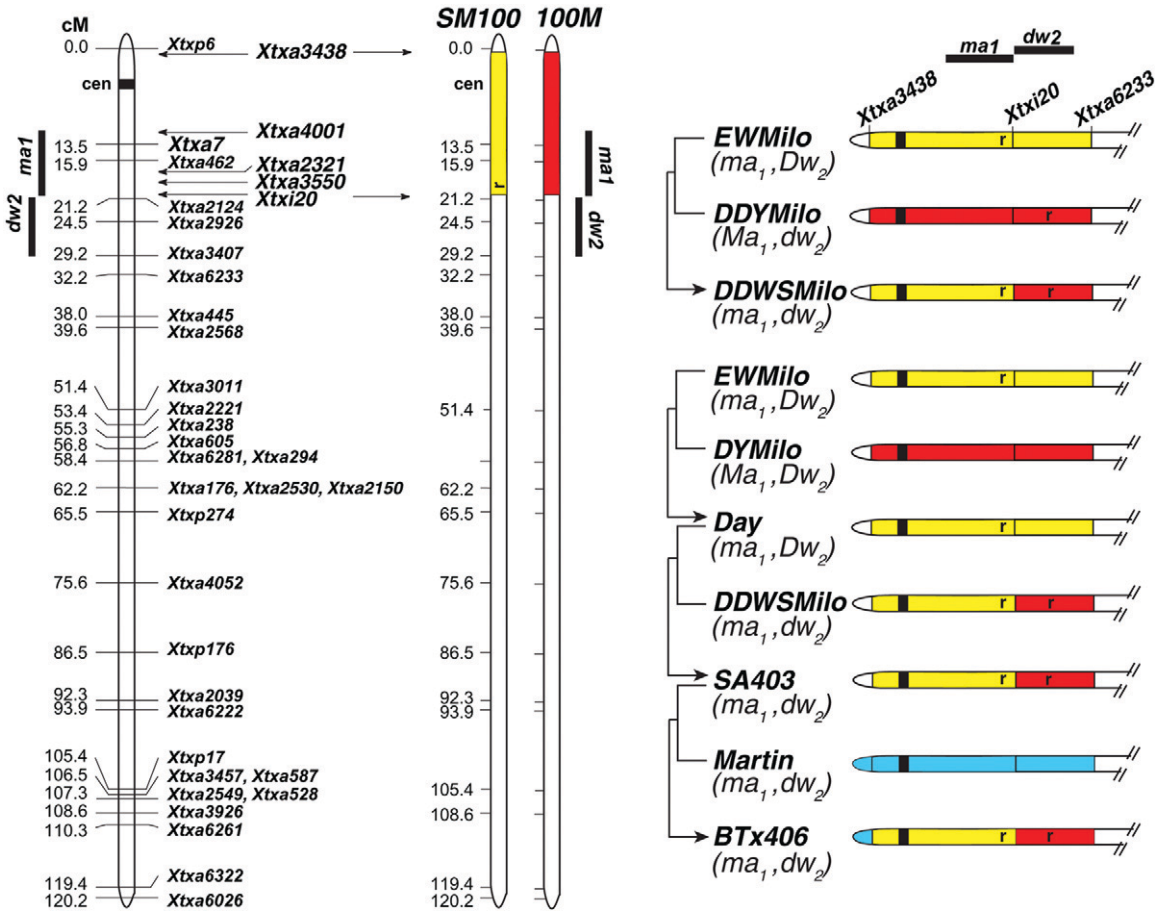


Figure 1. Map of sorghum linkage group SBI-06 depicting map location of the $ma_1$ and $dw_2$ loci. Left panel: Framework linkage map of SBI-06 displaying select framework markers plus the present resolution of the $ma_1$ and $dw_2$ loci (delimited with black bars). Center panel: Graphical representation of SBI-06 of SM100 and 100M depicting the single haplotype block that differentiates these two genetic stocks. Markers delineating this region are shown to the left while "r" depicts the block that harbors $ma_1$. Right panel: Graphical depiction of the inheritance of $ma_1$ and $dw_2$ genomic regions through the pedigree of BTx406. Haplotype blocks of similar color are identical by descent (IBD), while "r" represents blocks harboring the recessive allele of $ma_1$ or $dw_2$. The $ma_1$ and $dw_2$ genotype of each line are given to the left of the graphical genotypes. Cultivar abbreviations are shown in Supplemental Table 1.

(1974) observation that the linkage between $Dw_2$ and $ma_1$ can be readily broken.

## Graphical Genotypes of Founder and Early-Derived Cultivars

The graphical genotypes of SBI-06 for Milo and Kafir founder introductions are shown in Fig. 2A. The graphical genotypes of Standard Yellow Milo and Early White Milo revealed a high degree of identity, with only a single block of SBI-06 differentiating the two haplotypes. This divergent block (~21–32 cM) of SBI-06 abuts the haplotype block that differentiates SM100 and 100M (Fig. 1). Early White Milo ($ma_1$, $Ma_2$, $Ma_3$, $Ma_4$) originated as either a mutation in $Ma_1$ or as a selection from a seed blend of standard and early-maturing Milo genotypes. Standard Yellow Milo ($Ma_1$, $ma_2$, $ma_3$, $Ma_4$) originated by a mutation in $Ma_2$ and $Ma_3$, but retains the dominant allele at the $Ma_1$ locus. While day neutral in floral initiation, both Standard Yellow Milo ($Dw_1$, $Dw_2$, $Dw_3$, $dw_4$) and Early White Milo ($Dw_1$, $Dw_2$, $Dw_3$, $dw_4$) are tall (2–3 m). The dwarf cultivars, Dwarf Yellow Milo and Double Dwarf Yellow Milo, share a common haplotype with Standard Yellow Milo except for a single haplotype block at the end of SBI-06 proximal to $dw_2/ma_1$. Dwarf Yellow Milo ($dw_1$, $Dw_2$, $Dw_3$, $dw_4$) is believed to be a short-stature selection (circa 1905) from Standard Yellow Milo, resulting from a recessive allele at the $dw_1$ locus. Double Dwarf Yellow Milo was a short-stature selection from Dwarf Yellow Milo and differed from Dwarf Yellow Milo (and Standard Yellow Milo, Early White Milo) in harboring a dwarfing gene at the $Dw_2$ locus.

Examination of the graphical genotype of founder cultivar Blackhull Kafir reveals a haplotype that is unique from Milo genotypes (Fig. 2A). Blackhull Kafir

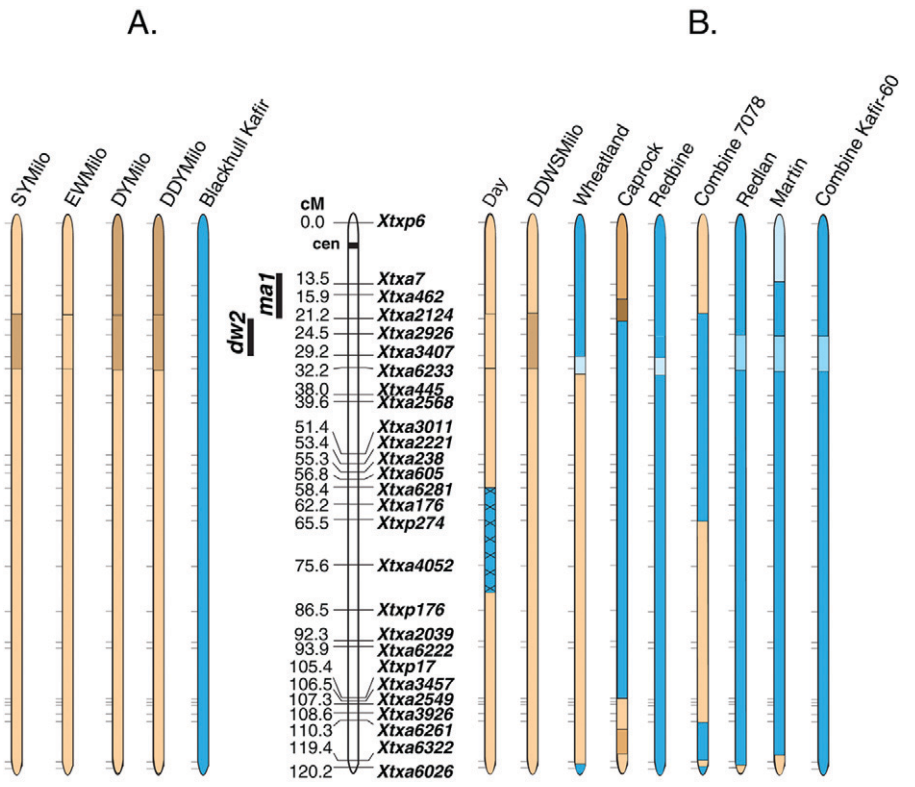


Figure 2. Graphical genotypes of founder introductions and early cultivar derivatives. The logical functions of Pedigree Based Marker Assisted Selection System (PBMASS) were utilized for haplotype block partitioning with flanking marker inference settings of 3 to 5 cM. Haplotype blocks of Milo ancestry are depicted in shades of tan–brown and orange, while blocks of Kafir descent are represented in shades of blue. Haplotype blocks of different cultivars that are either identical by descent (IBD) or identical by state (IBS) are depicted in the same color. Variants of ancestral haplotype blocks (e.g., differing in 1–2 marker score across the block) are represented in slightly different shades of the same base color. The cross-hatched block represents a region of Day Milo that was not consistent with the pedigree of the given cultivar. (A) Founder genotypes and derivatives selected for short stature and early maturation. (B) Popular public grain sorghum cultivars released from the 1920s to 1950s. Cultivar abbreviations are shown in Supplemental Table 1.

is an early maturing cultivar based on a recessive allele at the $ma_1$ locus (Quinby, 1974, 1975). Despite Blackhull Kafir and Early White Milo both encoding a recessive form of $ma_1$, the haplotype blocks spanning this locus are unique in the Kafir and Milo backgrounds. Quinby (1974) suggested that floral initiation in temperate climates was controlled by an allelic series at several maturity loci, including $ma_1$. The unique nature of the haplotype blocks spanning $ma_1$ in Early White Milo and Blackhull Kafir provides further support for an allelic series for the $ma_1$ locus.

## Graphical Genotypes of Cultivars from Milo and Kafir Crosses

From the original temperate Milo and Kafir cultivars were derived a series of important genotypes, primarily as natural or controlled crosses (Fig. 2B). Crosses of Early White Milo and Dwarf Yellow Milo resulted in dwarf, early-maturing cultivars including Day Milo (circa 1924) and Double Dwarf White Sooner Milo (circa 1930). The common pedigree of these cultivars is reflected in their graphical genotypes of SBI-06, which are IBD in extended regions of the chromosome. There are, however, haplotype blocks that are not shared by Day Milo and Double Dwarf White Sooner Milo, including a block extending from ~21–30 cM. Day Milo derived this haplotype block from Early White Milo, whereas the haplotype block of Double Dwarf White Sooner Milo is derived from the other parent, Dwarf Yellow Milo. Day Milo also exhibits a haplotype block that is IBS with Kafirs (blue, hatched block), which is noteworthy since Kafir is not reported in the pedigree of Day Milo. While the occurrence of marker haplotypes that are inconsistent with pedigrees was surprisingly uncommon, it did occur in a number of lines, particularly in the converted materials.

The pedigree of Wheatland, a selection from an early Kafir × Milo cross (circa 1931), is also reflected in the graphical genotype of SBI-06. The graphical genotype of Wheatland shows extended, uninterrupted haplotype blocks derived from Kafir (0–36 cM) or of Milo parentage (36–119 cM). Cultivars

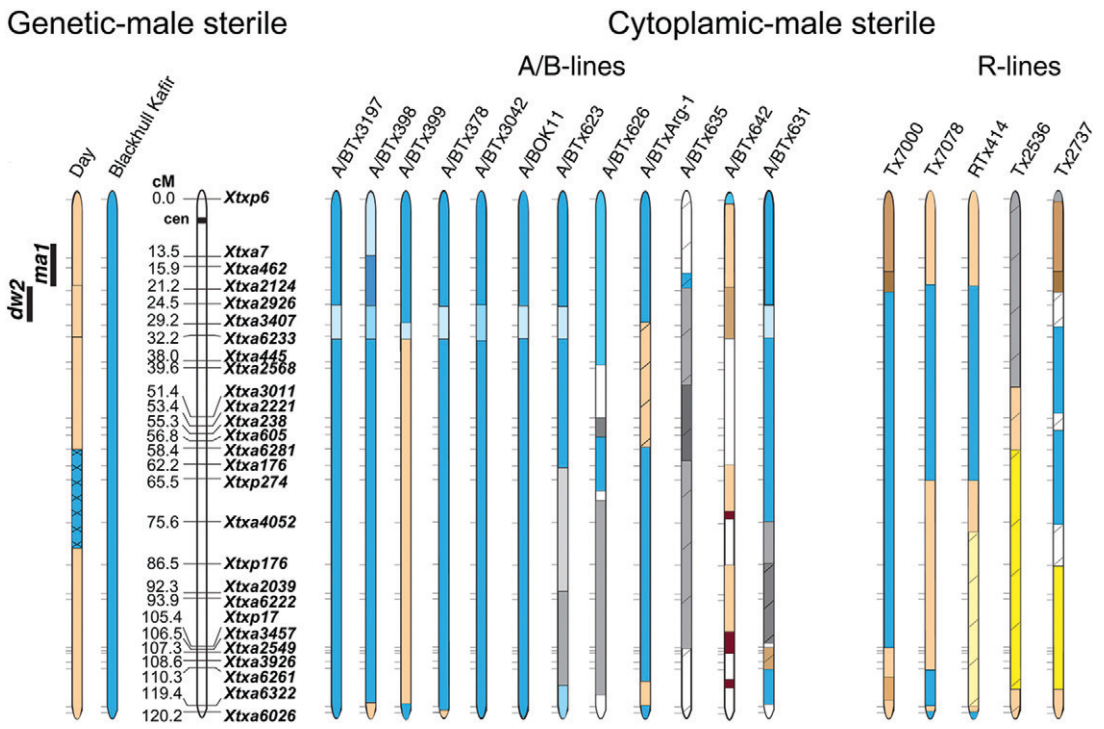


Figure 3. Graphical genotypes of early A- and R-lines central to the development of hybrid sorghum. Left panel: Genetic male sterile Day Milo (female) and pollinator Blackhull Kafir from early hybrid crosses. Right panel: Early generation cytoplasmic male sterility (CMS) A-lines (female) and R-lines (fertility restorer). Haplotype block partitioning and color designation is as described in Fig. 2, plus yellow-colored blocks depict regions of Feterita descent, gray-colored blocks are of Caudatum–Zerazera descent, and maroon blocks are of Durra descent. Uncolored blocks denote extended strings of uninformative monomorphic markers within ancestors of a pedigree, or blocks where IBS or IBD was not predicted by Pedigree Based Marker Assisted Selection System (PBMASS) or Excel's advanced filter tool. Blocks that are marked with diagonal lines denote blocks that are IBS with similarly colored genotypes, whereas cross-hatched blocks represent regions that were not consistent with the pedigree of the given cultivar.

developed for mechanical harvest including Caprock (Tx7000), Redbine (Tx3042), Combine 7078 (Tx7078), Redlan (BTx378), and Martin (BTx398), each show unique haplotypes, but also share extensive regions that are IBD. Martin and Redlan, whose phenotypes strongly resemble Kafir and not Milo, exhibit a strikingly Kafir-like haplotype with little evidence of the Milo parental haplotype. Martin and Redlan haplotypes are similar to Combine Kafir-60 (Tx3197), which was developed from a Kafir × Kafir cross. By contrast, Caprock and Combine 7078 more closely resemble Milo in appearance, which is consistent with the presence of haplotype blocks that are derived from the non-Kafir parent.

## Genotypes of Hybrid Sorghum's Origin

The graphical genotypes of cultivars central to the development of hybrid sorghum are shown in Fig. 3. Before the discovery of CMS in sorghum, genetic male sterility of Day Milo was used to produce hybrid seed (Fig. 3, left panel). Day Milo and Blackhull Kafir appeared on opposite sides of these first generation hybrids that yielded 40% more grain than the cultivars in general use in the mid-1940s. Consistent with the pedigrees of Day Milo and Blackhull Kafir, the haplotypes of SBI-06 are unique, with only one Kafir-like haplotype block being IBS between the two cultivars. The first generation CMS A-lines were sterilized Kafir cultivars, with Combine Kafir-60 (A/BTx3197) being the most widely used female in hybrid seed production (Fig. 3, right panel). Graphical genotypes reflect the prominence of Kafir in the pedigrees of A-lines with extended blocks being of Kafir descent. Several exceptions included A/BTx642 and A/BTx635. A/BTx635, which is a derivative of a photoperiod insensitive Zerazera cultivar, originated from ICRSAT (Miller et al., 1982), while ATx642 is a converted exotic line of Durra descent.

First generation R-lines were also discovered in popular Kafir-Milo cultivars including Caprock (Tx7000) and Combine 7078 (Tx7078). Genetic similarity estimates (Menz et al., 2004) would suggest that many of the first generation R- and A/B-lines would display considerable portions of the genome that are IBD. As predicted, an extended region of the SBI-06 haplotype of Tx7000 was of Kafir

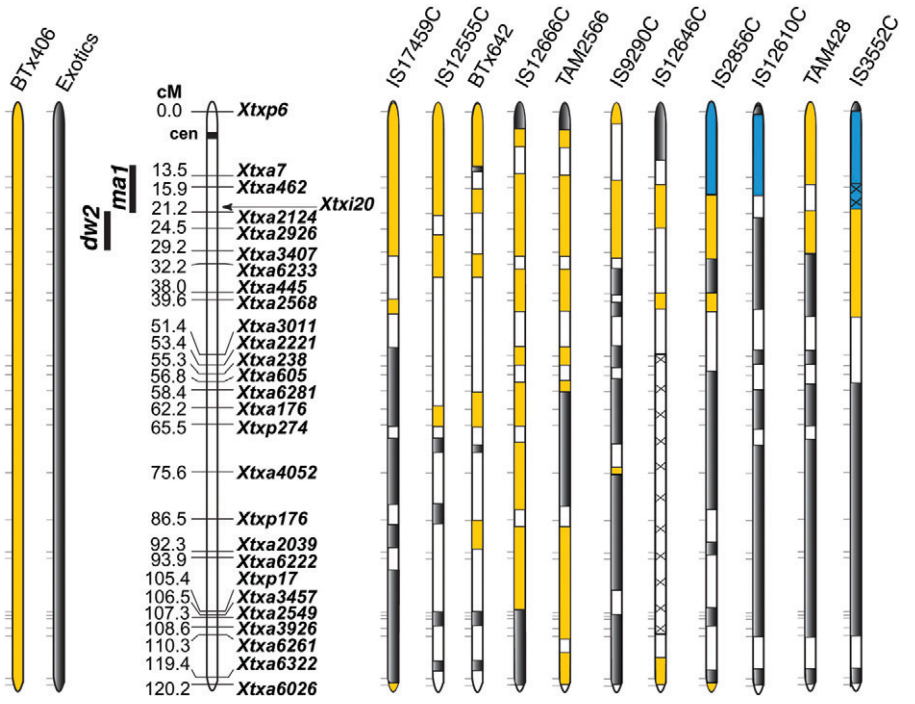


Figure 4. Introgression of donor haplotype blocks into the genome of tropical genotypes during their conversion to photoperiod insensitive, short-statured cultivars. Donor line (normally BTx406) haplotype blocks are depicted in orange, black shaded blocks represent blocks identical by descent (IBD) with the exotic parent, and uncolored regions of SBI-06 represent regions of the genome where markers were either missing or were not polymorphic between the donor and exotic line. Blue blocks (without cross-hatch) are derived from the exotic parent, but are identical by state (IBS) with Kafirs. Cross-hatched blocks represent regions that are not derivatives of the donor or the exotic parent.

descent. Tx7078 and RTx414 (a Tx7078 derivative) also displayed extended Kafir blocks, but additional regions of SBI-06 were of Milo descent (tan/brown) or were unique (light-yellow block of RTx414) being unlike any other genotype examined. Other early R-lines with excellent combining ability included Feteritas such as Tx09 and Combine White Feterita R3067 (data not shown). Preliminary marker analyses of these cultivars revealed haplotypes that are unique from any genotype examined in the present study (R.R. Klein, unpublished results, 2007). With the subsequent release of Tx2536 (Feterita and Caudatum Kaura pedigree, circa 1964) and Tx2737 (Tx7000–Tx2536 derivative, circa 1976 [Johnson et al., 1982]), additional male lines with notable diversity with A-lines were available, with this genetic diversity being reflected in the graphical genotypes of SBI-06.

## Tropical Sorghum Conversion and Introgression of Temperate Donor Genome

The breeding scheme used in the Sorghum Conversion Program involved the introgression of recessive $ma_1$ and $dw_2$ alleles from a temperate cultivar into exotic germplasm (Quinby, 1974; Rosenow and Dahlberg, 2000; Stephens et al., 1967). Figure 4 depicts regions of SBI-06 that were introgressed from the temperate donor line (most commonly, BTx406) during the conversion of select exotic lines. Displayed are regions of SBI-06 that were IBD with either the donor line (colored orange) or the exotic (shaded black), and regions where IBD cannot be determined (uncolored blocks) since the parental alleles were identical. The haplotype of each converted line ("C" suffix) is a mosaic of blocks of markers across the entire linkage group, showing extended regions that are IBD with the exotic parent or the temperate $ma_1$ donor line. Consistent with the conversion of exotic lines at the $ma_1$ locus, much of the region proximal to the mapped location of $ma_1$ (*Xtxi20*) were IBD with the donor parent for most converted lines. The notable exceptions were converted lines IS2856C, IS12610C, and IS3552C, which shared an extended exotic-derived marker block near $ma_1$ (blue designation). This exotic-derived haplotype block is IBS with early maturing Kafir cultivars, and may possess an alternate $ma_1$ allele.

Extending further from the mapped location of $ma_1$ and $dw_2$, no donor-derived region of SBI-06 was consistently introgressed into all converted lines. Exotic-derived regions of the converted lines were

more prevalent distal to $ma_1$ and $dw_2$, with IS12610C showing nearly complete recovery of the exotic haplotype. By contrast, the marker haplotypes of several converted lines (e.g., IS12666C) were extensively IBD with the donor line. The results further indicate that the extent to which the exotic haplotype was recovered was not strictly a function of the number of backcross generations. TAM2566, an early backcross release ($BC_2$) of exotic IS12666, showed a similar proportion of the exotic haplotype as IS12666C ($BC_4$). BTx642 ($BC_1$) and IS12555C ($BC_4$), two derivatives of IS12555, also showed similar SBI-06 haplotypes with extended blocks from the nonrecurrent donor line.

## Tropical Conversion and Haplotype Diversity of Sorghum Cultivars

Graphical genotypes of a set of public cultivars that figured prominently in sorghum improvement efforts during the 20th century are shown in Fig. 5. The conversion of tropical materials, the introduction of photoperiod insensitive cultivars, and the intermating of this germplasm dramatically increased the haplotype diversity of sorghum cultivars. As evidenced by the graphical genotypes and genetic similarity estimates, the incorporation of converted and temperate lines into the pedigrees of modern parental lines and cultivars effectively eliminated the bottlenecks created during early germplasm improvement efforts in the United States. Since most of the introduced and converted materials were able to restore pollen fertility, the greatest impact on haplotype diversity was observed in R-lines. By contrast, the conversion of tropical accessions had less impact on the haplotype diversity of A-lines since most females developed during the conversion era arose from crosses of early A-lines and partially converted tropical cultivars (e.g., A/BTx623, A/BTx626, A/BTx631). With the exception of A/BTx642, A-lines were not directly derived from exotic material, which limited the genetic diversity of this group.

Examination of the graphical genotypes of many converted lines and temperate introductions revealed that the region flanking the $ma_1$ locus was IBD (or IBS) with early-maturing Milo or Kafir. Exceptions to this included the photoperiod insensitive line Tx2536 and derivatives of this line (RTx430, Tx2908). The unique haplotype block spanning $ma_1$ in the photoperiod insensitive lines is not inconsistent with the existence of an extensive allelic series for this maturity gene. In examining the haplotype diversity extending further from the $ma_1$ region in converted cultivars, greater haplotype diversity was observed, which is consistent with the breeding objective of the conversion program.

As clear heterotic groups have not been created for public sorghum cultivars, graphical genotypes in Fig. 5 were clustered based on the genetic similarity estimates of Menz et al. (2004). In general, lines with high genetic similarity estimates (or those within a working group) showed extended regions of SBI-06 that were IBD or IBS. One notable exception to this was BTx406, which clustered with Kafirs, but displayed a Milo-like SBI-06 haplotype. The SBI-06 haplotypes of some cultivars within a common working group diverged in a very limited number of marker blocks; in several cases (e.g., Tx2536 and RTx430), a slight variant of a single haplotype block can differentiate SBI-06 in closely related cultivars. Several lines show haplotypes that were quite unique, including IS3620C and IS6705C. The unique graphical genotype of these two cultivars was reflected in the genetic similarity estimates of these converted lines.

The present results provide an opportunity to examine the relationship between graphical genotypes and genetic similarity estimates in the context of the known pedigrees of sorghum cultivars. While a detailed analysis is beyond the scope of the present study, several interesting trends are apparent. In examining Tx2908, RTx430, and Tx2536, the shared parentage of these lines is reflected in graphical genotypes and in genetic similarity estimates. RTx430 is a parental line of Tx2908, and Tx2536 is in the pedigree of RTx430. Large haplotype blocks of RTx430, Tx2908, and Tx2536 are IBD or IBS, which is consistent with genetic similarity estimates. In contrast, genetic similarity estimates indicate a close relationship between RTx432 and BTx623, but this predicted relationship was not reflected in the graphical genotypes of SBI-06. It should be noted that RTx432 and BTx623 do not share a common pedigree, which raises concerns with genetic similarity estimates for these two cultivars. Further examination of genetic similarity estimates and the graphical genotypes of the remaining sorghum chromosomes may lead to a better understanding of this discrepancy. Finally, RTx430 and BTx623 represent an interesting paradigm. RTx430 and BTx632 are half sibs sharing a common parent in IS12661C. Despite sharing a common parent, cluster analysis and graphical genotypes did not reflect this genetic relationship. Examination of graphical genotypes of RTx430 and BTx623 provides an explanation for this apparent contradiction. The graphical genotype of RTx430 resembles that of parental line Tx2536, with little similarity to IS12661C, while BTx623 shows extended haplotype blocks that are IBD with parental line BTx3197. Thus, while RTx430 and BTx623 are half

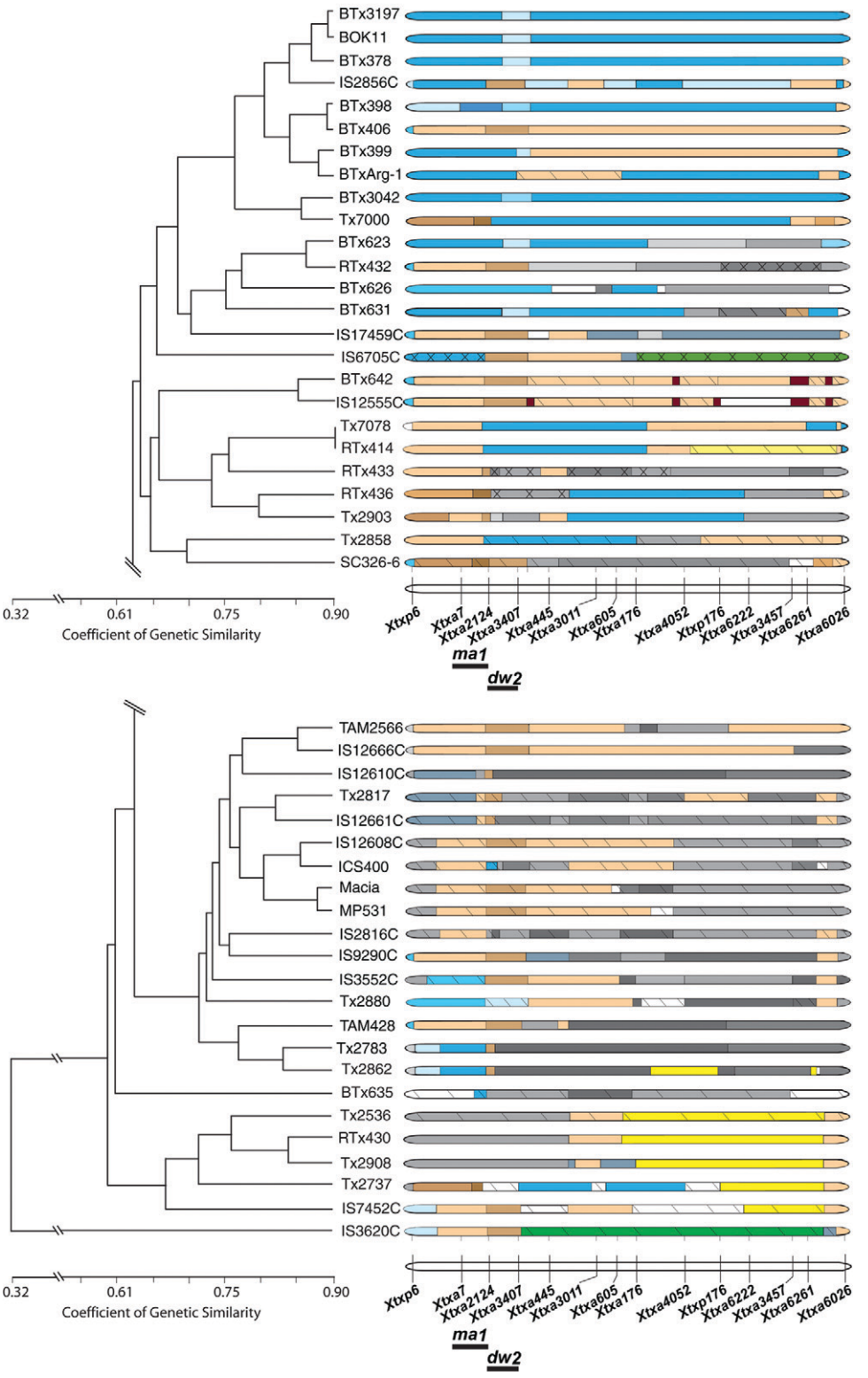


Figure 5. Graphical genotypes of a set of public cultivars that figured prominently in sorghum improvement during the 20th century. Cultivars are ordered based on genetic similarity estimates as previously determined (Menz et al., 2004). Haplotype blocks are designated as detailed for Fig. 2 to 4, plus green representing blocks of Guinea descent. Variants of haplotype blocks are depicted as different shades of the same base color. New base colors for a haplotype block are used as new marker sequences are identified. The fertility reaction of the lines (e.g., A- and R-line designation), is shown in Supplemental Table 1.

sibs, genetic similarity estimates and the graphical genotypes of BTx623 and RTx430 do not reflect this relationship.

## Discussion

Identification and characterization of the genetic variation underlying traits of agronomic importance is central to all crop improvement programs. There now exists an opportunity to detail the genetic diversity for defined chromosomal regions by constructing high-resolution graphical genotypes for any number of selected cultivars. Graphical genotypes can be viewed as a molecular form of pedigree analysis allowing breeders to retrospectively evaluate the genomic composition of each cultivar and to determine the contribution of each progenitor to specific genomic regions (Jordan et al., 2004). To this end, we examined the phylogenetic relationship of sorghum cultivars within segmental blocks of the sorghum genome enabling us to more precisely characterize the patterns of genetic variation among classical inbreds and determine the relative contributions of cultivar-specific and shared ancestral variation in the sorghum genome. The graphical genotypes of landmark cultivars are expected to reveal characteristic signatures of crop improvement efforts such as man-made bottlenecks, migration, nonrandom mating, and directional selection (Hamblin et al., 2004, 2005, 2006; Jones et al., 1995; Wade et al., 2002; Wade and Daly, 2005; Yalcin et al., 2004). By examining cultivars released throughout the 20th century, genomic signatures of historical breeding decisions will be apparent with the resolution of these events being limited largely by the number of cultivars examined and the density of markers within a specified chromosomal region.

It is recognized that because of a man-made bottleneck created during sorghum's introduction into the United States, the genomes of first generation cultivars originated from a mixed, but very limited pool of founders. Of the grain sorghum cultivars of the first 40 years of the 20th century, 75 to 80% are derivatives of the original Milo and Kafir introductions (Smith and Frederiksen, 2000). Consistent with the bottleneck created during sorghum's introduction, the chromosome SBI-06 graphical genotypes of first generation sorghum cultivars display extended blocks of the genome that are either of Kafir or Milo ancestry. Many of the haplotype blocks were extensive, covering vast expanses of SBI-06, which is consistent with the highly inbred nature of sorghum and the breeding history before the establishment of the Sorghum Conversion Program. Also included among the introductions important to the first 40 years of sorghum production were the Feterita types (1908) from the Sudan of Africa. The graphical genotypes of cultivars with Feterita ancestry (e.g., Tx2536, RTx430, Tx2908) reveal a haplotype that diverges substantially from Milo or Kafir, indicating the importance of this introduction in increasing the diversity of existing germplasm sources.

With the discovery of genetic male sterility and subsequently CMS in sorghum, the era of hybrid seed production began in 1956. The breeding history of sorghum before the production of hybrids involved the development of cultivars from the intermating of a limited number of plant introductions and their derivatives. Armed with the extended pedigrees of cultivars, graphical genotypes were used to follow allele flow through every step in the development of lines that would impact early efforts in hybrid sorghums. It is clear that early breeding methods, especially the practice of intermating cultivars, compromised the potential heterotic patterns of sorghum. By intermating cultivars of different sorghum working groups, many public cultivars display large haplotype blocks of common ancestry. Their haplotypes are therefore admixtures of the genomes of two potential heterotic groups. Selection pressure and genetic hitchhiking in this highly inbred crop has led to these common haplotype blocks being rather expansive in some cases. The impact of these breeding practices on the development of suitable female and male lines was revealed with the discovery of CMS in sorghum. As male-sterile cytoplasm was discovered in Kafir, existing Kafirs (e.g., Combine Kafir-60) and Kafir-Milo derivatives (e.g., BTx398, BTx399) were sterilized to produce suitable A-lines. Many of the early R-lines including Tx7000 and Tx7078 were also of Kafir or Kafir–Milo descent. As expected, SBI-06 graphical genotyping identified extensive regions of the genome of A- and R-lines that are IBD (A- and R-line designation are shown in the supplemental material). Breeders have recognized that the shared ancestry of A- and R-lines had in fact compromised the potential heterotic pattern. Fortunately, additional R-lines with excellent combining ability were developed from races including Hegari and Feterita. The haplotype of Tx2536, a cultivar of Caudatum and Feterita pedigree, shared little identity with the haplotypes of all other early introductions. This diversity, reflected in the divergence of A- and R-line haplotypes, was central to developing lines with high combining ability.

With the advent of the hybrid era and the man-made bottleneck created during sorghum's introduction, new diverse germplasm with useful genes and traits were sought by sorghum breeders. Sorghum cultivars in the world collection number in the thousands, but most are photoperiod sensitive

and mature too late for temperate-zone production. Armed with knowledge of the inheritance of plant height and maturity in sorghum, the USDA-TAES Sorghum Conversion Program was established in 1963 to convert many of the tropical accessions to temperate adaptation (Stephens et al., 1967). The program has been very effective in providing new, diverse germplasm, with most of the best and currently used sources of resistance to diseases and abiotic stresses being derived from converted exotics (Rosenow and Dahlberg, 2000). The scheme was designed to move recessive dwarfing and maturity genes from a four-dwarf temperate zone variety into an exotic background. Based on the conventional wisdom, fully converted sorghum genotypes were expected, in theory, to be comprised of 97% recurrent (exotic-tropical) parent genome; thus, most of the remaining donor chromatin was to contain alleles for reduced height and photoperiod-insensitive flowering (Lin et al., 1995). The graphical genotypes of converted tropical accessions demonstrate the resulting divergence of genomic regions in both the genome of converted accessions and derived cultivars. However, recovery of the exotic genome of SBI-06 was not nearly as complete as was suggested (Lin et al., 1995), with extensive stretches of the donor genome surviving through four backcross generations. The persistence of donor regions in converted genomes distal to the height or maturity genes was not the result of intentional selection pressure, but may reflect genetic hitchhiking and background selection caused by the elimination of deleterious alleles (Hamblin et al., 2004, 2005, 2006). Various regions of the genome will encode genes that will condition agronomic fitness of derived backcross lines including tolerance to biotic and abiotic stresses. Culling of weak or diseased backcrossed individuals would perform background selection against the tropical genome if this genomic region of the donor line conferred fitness in backcross-derived lines. Further understanding of natural and breeder-induced selection's impact on the genome of sorghum cultivars is purely speculative at this point, but clearly, recovery of 97% of the exotic genome was not achieved during the conversion process.

There now exists an opportunity to detail the genetic diversity for defined chromosomal regions by constructing high-resolution graphical genotypes for any number of selected cultivars.

Consistent with the conversion of exotic lines at the $ma_1$ locus, the regions proximal to the mapped location of $ma_1$ (*Xtxi20*) were IBD with the donor parent for many converted lines. Similar results have been reported previously by Lin et al. (1995). However, following the bottleneck created by a limited number of founder genotypes, the signal of directional selection may be hard to detect in early cultivars because many loci may have low variation by chance and this may resemble the effects of selection. The effect of this bottleneck, however, was effectively minimized by the incorporation of exotic germplasm into the gene pool via the Sorghum Conversion Program. Hence, the observed reduction in diversity proximal to $ma_1/dw_2$ should reflect, to a considerable extent, the effect of directional selection during the temperate adaptation of sorghum cultivars. The extent of introgression of the donor haplotype flanking the $ma_1/dw_2$ genomic region also suggests that genetic hitchhiking persisted during directional selection for short-stature, early-maturing phenotypes. This was not unexpected since hitchhiking effects of directional selection will be prevalent in self-pollinating species due to reduced effective rate of recombination (Hamblin et al., 2004). The $ma_1/dw_2$ loci of the donor line BTx406 were of Milo descent and hence, most all converted lines will possess a Milo-like haplotype spanning $ma_1$ and $dw_2$. Other temperate-adapted cultivars, including several converted lines, displayed a Kafir- or Feterita-like haplotype block spanning this region of SBI-06. Hence, the graphical genotypes presented herein provide confirmatory evidence for the existence of an allelic series at the $ma_1$ locus, with mutations having occurred independently in the genetic backgrounds of the different sorghum races. As the genomes of more temperate accessions are examined, it is likely that the size of the allelic series for $ma_1$ will expand, and more unique haplotype blocks tagging the locus will be revealed, and thereby further complicating the use of markers to screen for $^{r}ma_1$ during tropical germplasm conversion.

The apparent mosaic structure of the sorghum genome is being exploited by our group for quantitative trait loci (QTL) mapping in several ways. First, it focuses the search for functional variants into divergent haplotype blocks. By examining a haplotype map describing segments of different ancestry, we can identify a smaller subset of the region in which the haplotypes diverge in accordance with the phenotype. Knowledge of the patterns of variation among phenotyped inbred strains will do much to reduce the search space for detecting QTL with the available

genetic resources. Using comprehensive catalogs of haplotype variation, in conjunction with extensive characterizations of cultivar phenotypes and expression patterns, will become a routine and enabling approach to positional cloning in sorghum. Graphical genotypes will also facilitate the identification of appropriate parental genotypes for creation of populations for confirmatory mapping of QTL. Selection of parental lines for linkage analysis will therefore be based not only on phenotypic differences, but also on haplotype divergence in the targeted region. Having complete phenotypic and genotypic data across numerous cultivars, recombinant inbred lines, and other germplasm may provide a wealth of information regarding the architecture of complex phenotypes.

In the future, unlocking the full potential of the sorghum genome for genetic analysis will require a dense sequenced-based genetic map and an understanding of the structure of genetic variation both within and among cultivars. Our present marker density provides a rather coarse picture, and greater marker density will allow further resolution of regions that diverge across loci of interest. As linkage disequilibrium estimates among sorghum elite cultivars are predictably high, a limited number of haplotype-tagging markers will allow tracing this region through germplasm panels. Population genetic studies can complement QTL studies by identifying regions that have been subjected to human selection. The genomic region affected by selection may be relatively large in sorghum due to the reduced effective recombination rate. Nevertheless, a haplotype map of an exhaustive set of public cultivars will offer a precise picture of the history of each individual genomic region, which is likely to become a valuable complement to traditional approaches in linkage mapping.

## Acknowledgments

The authors wish to thank Ms. Julie McCollum (USDA-ARS) and Natalie Unruh (Texas A&M University) for expert technical assistance. The research is supported by the United States Department of Agriculture's Agricultural Research Service (R.R.K.). This work was also supported in part by National Science Foundation Plant Genome Grant DBI-0321578 (P.E.K., R.R.K., and J.E.M.).

## References

Bowers, J.E., C. Abbey, S. Anderson, C. Chang, X. Draye, A.H. Hoppe, R. Jessup, C. Lemke, J. Lennington, Z.K. Li, Y.R. Lin, S.C. Liu, L.J. Luo, B.S. Marler, R. Ming, S.E. Mitchell, D. Qiang, K. Reischmann, S.R. Schulze, D.N. Skinner, Y.W. Wang, S. Kresovich, K.F. Schertz, and A.H. Paterson. 2003. A high-density genetic recombination map of sequence-tagged sites for *Sorghum*, as a framework for comparative structural and evolutionary genomics of tropical grains and grasses. Genetics 165:367–386.

Bowers, J.E., M.A. Arias, R. Asher, J.A. Avise, R.T. Ball, G.A. Brewer, R.W. Buss, A.H. Chen, T.M. Edwards, J.C. Estill, H.E. Exum, V.H. Goff, K.L. Herrick, C.L.J. Steele, S. Karunakaran, G.K. Lafayette, C. Lemke, B.S. Marler, S.L. Masters, J.M. McMillan, L.K. Nelson, G.A. Newsome, C.C. Nwakanma, R.N. Odeh, C.A. Phelps, E.A. Rarick, C.J. Rogers, S.P. Ryan, K.A. Slaughter, C.A. Soderlund, H.B. Tang, R.A. Wing, and A.H. Paterson. 2005. Comparative physical mapping links conservation of microsynteny to chromosome structure and recombination in grasses. Proc. Natl. Acad. Sci. USA 102:13206–13211.

Conner, A.B., and R.E. Karper. 1927. Hybrid vigor in sorghum. Texas Agric. Exp. Stn. Bull. 359. Texas A&M Univ., College Station.

Hamblin, M.T., A.M. Casa, H. Sun, S.C. Murray, A.H. Paterson, C.F. Aquadro, and S. Kresovich. 2006. Challenges of detecting directional selection after a bottleneck: Lessons from *Sorghum bicolor*. Genetics 173:953–964.

Hamblin, M.T., S.E. Mitchell, G.M. White, W. Gallego, R. Kukatla, R.A. Wing, A.H. Paterson, and S. Kresovich. 2004. Comparative population genetics of the panicoid grasses: Sequence polymorphism, linkage disequilibrium and selection in a diverse sample of *Sorghum bicolor*. Genetics 167:471–483.

Hamblin, M.T., M. Salas Fernandez, A.M. Casa, S.E. Mitchell, A.H. Paterson, and S. Kresovich. 2005. Equilibrium processes cannot explain high levels of short- and medium-range linkage disequilibrium in the domesticated grass *Sorghum bicolor*. Genetics 171:1247–1256.

Johnson, J.W., D.T. Rosenow, G.L. Teetes, and J.M. Phillips. 1982. Registration of 19 greenbug resistant sorghum germplasm lines. Crop Sci. 22:1272.

Jones, D.L., J.E. Shaff, and L.V. Kochian. 1995. Role of calcium and other ions in directing root hair tip growth in *Limnobium stoloniferum*: I. Inhibition of tip growth by aluminum. Planta 197:672–680.

Jordan, D.R., Y.Z. Tao, I.D. Godwin, R.G. Henzell, M. Cooper, and C.L. McIntyre. 2004. Comparison of identity by descent and identity by state for detecting genetic regions under selection in a sorghum pedigree breeding program. Mol. Breed. 14:441–454.

Klein, P.E., R.R. Klein, S.W. Cartinhour, P.E. Ulanch, J. Dong, J.A. Obert, D.T. Morishige, S.D. Schlueter, K.L. Childs, M. Ale, and J.E. Mullet. 2000. A high-throughput AFLP-based method for constructing integrated genetic and physical maps: Progress toward a sorghum genome map. Genome Res. 10:789–807.

Klein, P.E., R.R. Klein, J. Vrebalov, and J.E. Mullet. 2003. Sequence-based alignment of sorghum chromosome 3 and rice chromosome 1 reveals extensive conservation of gene order and one major chromosomal rearrangement. Plant J. 34:605–621.

Klein, R.R., R. Rodriguez-Herrera, J.A. Schlueter, P.E. Klein, Z.H. Yu, and W.L. Rooney. 2001. Identification of genomic regions that affect grain-mould incidence and other traits of agronomic importance in sorghum. Theor. Appl. Genet. 102:307–319.

Lin, Y.R., K.F. Schertz, and A.H. Paterson. 1995. Comparative analysis of QTLs affecting plant height and maturity across the *Poaceae*, in reference to an interspecific Sorghum population. Genetics 141:391–411.

Lorenzen, L.L., S. Boutin, N. Young, J.E. Specht, and R.C. Shoemaker. 1995. Soybean pedigree analysis using map-based molecular markers: I. Tracking RFLP markers in cultivars. Crop Sci. 35:1326–1336.

Menz, M.A., R.R. Klein, J.E. Mullet, J.A. Obert, N.C. Unruh, and P.E. Klein. 2002. A high-density genetic map of *Sorghum bicolor* (L.) Moench based on 2926 AFLP, RFLP and SSR markers. Plant Mol. Biol. 48:483–499.

Menz, M.A., R.R. Klein, N.C. Unruh, W.L. Rooney, P.E. Klein, and J.E. Mullet. 2004. Genetic diversity of public inbreds of sorghum determined by mapped AFLP and SSR markers. Crop Sci. 44:1236–1244.

Miller, F.R., R. Clara, and V. Guiragossian. 1982. Registration of A/BTx635 sorghum. Crop Sci. 22:1517–1518.

Pestsova, E., and M. Roder. 2002. Microsatellite analysis of wheat chromosome 2D allows the reconstruction of chromosomal inheritance in pedigrees of breeding programmes. Theor. Appl. Genet. 106:84–91.

Quinby, J.R. 1974. Sorghum improvement and the genetics of growth. Texas A&M Univ. Press, College Station.
Quinby, J.R. 1975. The genetics of sorghum improvement. J. Hered. 66:56–62.
Rooney, W.L., and S. Aydin. 1999. Genetic control of a photoperiod-sensitive response in *Sorghum bicolor* (L.) Moench. Crop Sci. 39:397–400.
Rooney, W.L., and C.W. Smith. 2000. Techniques for developing new cultivars. p. 329–347. *In* C.W. Smith and R.A. Frederiksen (ed.) Sorghum: Origin, history, technology and production. John Wiley & Sons, New York.
Rosenow, D.T., and J.A. Dahlberg. 2000. Collection, conversion and utilization of sorghum. p. 305–328. *In* C.W. Smith and R.A. Frederiksen (ed.) Sorghum: Origin, history, technology, and production. John Wiley & Sons, New York.
Russell, J.R., R.P. Ellis, W.T.B. Thomas, R. Waugh, J. Provan, A. Booth, J. Fuller, P. Lawrence, G. Young, and W. Powell. 2000. A retrospective analysis of spring barley germplasm development from "foundation genotypes" to currently successful cultivars. Mol. Breed. 6:553–568.
Shoemaker, R.C., R.D. Guffy, L.L. Lorenzen, and J.E. Specht. 1992. Molecular genetic mapping of soybean: Map utilization. Crop Sci. 32:1091–1098.
Sjakste, T.G., I. Rashall, and M.S. Roder. 2003. Inheritance of microsatellite alleles in pedigrees of Latvian barley varieties and related European ancestors. Theor. Appl. Genet. 106:539–549.
Smith, C.W., and R.A. Frederiksen. 2000. History of cultivar development in the United States: From "Memoirs of A.B. Maunder—Sorghum breeder." p. 191–223. *In* C.W. Smith and R.A. Frederiksen (ed.) Sorghum: Origin, history, technology, and production. John Wiley & Sons, New York.
Stephens, J.C., and R.F. Holland. 1954. Cytoplasmic male sterility for hybrid sorghum seed production. Agron. J. 46:20–23.
Stephens, J.C., F.R. Miller, and D.T. Rosenow. 1967. Conversion of alien sorghum to early combine types. Crop Sci. 7:396.
Ulanch, P.E. 1999. Steps towards map-based cloning Ma(1), a gene controlling photoperiod sensitivity in sorghum. Ph.D. thesis, Texas A&M Univ., College Station.
Wade, C.M., and M.J. Daly. 2005. Genetic variation in laboratory mice. Nat. Genet. 37:1175–1180.
Wade, C.M., E.J. Kulbokas, A.W. Kirby, M.C. Zody, J.C. Mullikin, E.S. Lander, K. Lindblad-Toh, and M.J. Daly. 2002. The mosaic structure of variation in the laboratory mouse genome. Nature 420:574–578.
Yalcin, B., J. Fullerton, S. Miller, D.A. Keays, S. Brady, A. Bhomra, A. Jefferson, E. Volpi, R.R. Copley, J. Flint, and R. Mott. 2004. Unexpected complexity in the haplotypes of commonly used inbred strains of laboratory mice. Proc. Natl. Acad. Sci. USA 101:9734–9739.